

ABSTRACT

An automatic file clustering algorithm enables documents within a file system to be displayed in a semantic view. The file clustering algorithm maps all words and documents into an appropriate semantic vector space, clusters the documents at a predetermined level of granularity, and assigns a meaningful descriptor to each resulting cluster. The documents are displayed to the user in a hierarchy in accordance with the resulting clusters. This results in a virtual file system with a semantic organization, that allows the user to navigate by content.